Refined Composite Multivariate Generalized Multiscale Fuzzy Entropy: A Tool for Complexity Analysis of Multichannel Signals

Refined Composite Multivariate Generalized Multiscale Fuzzy Entropy: A Tool for Complexity Analysis of Multichannel Signals

Hamed Azami[1] and Javier Escudero[2]

[1,2]*Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh, King's Buildings, EH9 3JL, UK*

Emails:[1]hamed.azami@ed.ac.uk; [2]javier.escudero@ed.ac.uk

Abstract: Multiscale entropy (MSE) is an appealing tool to characterize the complexity of time series over multiple temporal scales. Recent developments in the field have tried to extend the MSE technique in different ways. Building on these trends, we propose the so-called refined composite multivariate multiscale fuzzy entropy (RCmvMFE) whose coarse-graining step uses variance (RCmvMFE$_{\sigma^2}$) or mean (RCmvMFE$_{\mu}$). We investigate the behavior of these multivariate methods on multichannel white Gaussian and 1/*f* noise signals, and two publicly available biomedical recordings. Our simulations demonstrate that RCmvMFE$_{\sigma^2}$ and RCmvMFE$_{\mu}$ lead to more stable results and are less sensitive to the signals' length in comparison with the other existing multivariate multiscale entropy-based methods. The classification results also show that using both the variance and mean in the coarse-graining step offer complexity profiles with complementary information for biomedical signal analysis. We also made freely available all the Matlab codes used in this paper.

Keywords: Complexity, multivariate generalized multiscale entropy, statistical moments, fuzzy entropy, sample entropy, biomedical signals.

1. Introduction

Entropy is a prevalent method to quantify the regularity of physical systems and to compare time series. To quantify the degree of the irregularity, randomness, or unpredictability of signals, a number of entropy measures were introduced during the past few decades [1-5]. One of the most popular kinds of entropy methods is sample entropy (SampEn) that measures the degree of randomness or, inversely, the degree of orderliness of a signal [2]. Since SampEn is less sensitive to the signal length and noise than approximate entropy (ApEn), it has been broadly used in biomedical signal processing [2]. Despite its popular use, SampEn is very sensitive to the threshold value. To tackle this problem, fuzzy entropy (FuzEn) was proposed [3]. These two entropy methods have attracted a great deal of attention over the recent years [6-11].

SampEn and FuzEn approaches, though powerful, are estimated only at a single temporal scale and therefore, may fail to account for the multiple time scales underlying nonlinear dynamics [12]. As an example, although the SampEn value of white Gaussian noise (WGN) signal is higher than that of $1/f$ noise, $1/f$ noise is theoretically more complex than WGN because of the long-range correlations of the former [13]. To overcome this shortcoming, multiscale entropy (MSE) [13, 14] and multiscale FuzEn (MFE) [3, 15] were proposed to take into account the various scales of a signal. It is worth noting that, in this context, the "complexity" concept stands for "meaningful structural richness", which may be in contrast with regularity measures defined from classical entropy approaches such as ApEn, SampEn, and FuzEn. For example, ApEn was proposed to quantify the degree of predictability of signals [16]. Thus, ApEn is primarily a

"regularity" statistic, not a direct index of physiological complexity. SampEn and FuzEn are based on the ApEn, leading to regularity measures [2, 3]. Thus, the entropy of 1/*f* noise is lower than that of WGN at scale factor 1 using MSE [13].

In fact, the least complexity illustrates either a completely ordered system with a small entropy value or a completely disordered system with maximum entropy value [13, 14, 16-18]. For instance, WGN is more irregular than 1/*f* noise although the latter is more complex, because 1/*f* noise contains long-range correlations and its 1/*f* decay produces a fractal structure in time. As another example, traditional entropy-based methods assign higher entropy values to certain pathologic cardiac rhythms that generate erratic outputs than to normal cardiac rhythms that are precisely regulated by multiple interacting control mechanisms [13, 14]. In the physiological complexity literature, healthy systems or people correspond to high complexity due to their ability to adapt themselves in response to adverse conditions, exhibiting long range correlations and complex variability at multiple scales, while aged and diseased systems or individuals present complexity loss, that is, they lose the capability to adapt to such adverse conditions [13, 16, 19].

In the MSE and MFE approaches, the original signal is initially divided into non-overlapping segments of length β, termed the scale factor. Next, the average of each segment is estimated to obtain the coarse-grained signals. Finally, the SampEn or FuzEn measure is calculated for each segment [13]. Costa and Goldberger have very recently generalized the MSE method using the second moment (variance) rather than the first moment (mean), in the coarse graining step of MSE [20]. This was named MSE_{σ}^2. It should be added that to discriminate MSE_{σ}^2 and the basic MSE, we will show the latter as MSE_{μ}.

MSE_{σ}^2 quantifies the dynamical properties of volatility (variance) over multiple time scales. MSE_{σ}^2 was applied to heartbeat signals from healthy young and older adults, and patients with congestive heart failure syndrome. The results showed that human heartbeat volatility signals depict complex bursting behaviours over a wide range of time scales. In addition, they found the multiscale complexity of both the volatility and mean degrades with aging [13, 20].

Multivariate signals, like multichannel recordings, are becoming more and more common in neuroscience and biomedical and mechanical science [21-23]. The MSE-based approaches, though powerful and widespread, are not able to reveal the dynamics across the channels. For such time series, evaluation of cross-statistical properties between multiple channels is essential for a complete understanding of the underlying signal-generating system [21, 24]. In this sense, Ahmed and Mandic proposed multivariate SampEn ($mvSE_{\mu}$) [21], leading to an extension of MSE to multivariate signals ($mvMSE_{\mu}$) [21]. The mvMSE analysis is interpreted based on 1) the multivariate signal **X** is more complex than the multivariate signal **Y**, if for the most time scales, the mvSE measures for time series **X** are larger than those for time series **Y**, 2) a monotonic fall in the multivariate entropy measures along the time scale factor demonstrates that the time series in hand only includes useful information at the smallest scales, and 3) a multivariate system illustrating long-range correlations and complex creating dynamics is characterized by either a constant mvSE or this declares a monotonic rise in mvSE with the time scale factor [21].

mvMSE has received much attention in the biomedical and mechanical fields [22, 23]. Labate and company employed the mvMSE and multivariate multiscale permutation entropy (mvMPE) [25] to predict the conversion from mild cognitive impairment to Alzheimer's disease using EEG signals [22]. Gao *et al.* proposed a multiscale complex network and multiscale clustering coefficient entropy to analyze multivariate time series [26]. They also used the mvMSE method

to characterize flow behavior underlying horizontal oil–water flows from experimental measurements [27].

Although mvSE is a powerful and popular algorithm, when applied to short time series, the results may be undefined or unreliable. To alleviate this limitation, we extend the refined composite MSE (RCMSE) [28], which was proposed for univariate signals, to multivariate time series. The mvMFE$_\mu$ method has been recently proposed to improve the stability of mvMSE$_\mu$ [15]. However, this approach, though powerful, is slow. In this paper, using another fuzzy membership function, the running time of mvMFE$_\mu$ is noticeably improved. Finally, we extend and investigate the new moment for coarse-graining process, variance, proposed for univariate signals, to multivariate signals. These methods are named as RCmvMFE$_\mu$ and RCmvMFE$_{\sigma^2}$.

This paper is organized as follows. In the next section, RCmvMFE$_{\sigma^2}$ and RCmvMFE$_\mu$ are presented in detail. In Section 3, the synthetic and real biomedical signals, employed in this piece of research, are described. The results and discussion are provided in Section 4. Finally, a conclusion is presented in Section 5.

2. Refined Composite Multivariate Multiscale Fuzzy Entropy

All multivariate multiscale sample/fuzzy entropy-based algorithms include two main steps as follows:

2.1. Coarse-graining for Multiscale Evaluation of Multivariate Entropy

A "coarse-graining" process is applied to a p-variate (channel) time series $\mathbf{Y} = \{y_{k,b}\}_{b=1}^{C}$, k=1,…,p, where C is the length of the signal at each channel. According to (1), each element of the coarse-grained signal is defined as:

$$ {}^{\mu}x_{k,i}^{(\beta)} = \frac{1}{\beta} \sum_{b=(i-1)\beta+1}^{i\beta} y_{k,b} \qquad 1 \le i \le \left\lfloor \frac{C}{\beta} \right\rfloor = N, \;\; 1 \le k \le p \tag{1} $$

where β is the time scale factor [14, 29]. Costa *et al.* [20] has recently proposed to use the variance, instead of mean value, in coarse-graining step in univariate signals. As an extension of this technique, we use variance in coarse-graining step for multi-channel time series as follows:

$$ {}^{\sigma^2}x_{k,i}^{(\beta)} = \frac{1}{\beta} \sum_{b=(i-1)\beta+1}^{i\beta} (y_{k,b} - {}^{\mu}x_{k,i}^{(\beta)})^2 \qquad 1 \le i \le \left\lfloor \frac{C}{\beta} \right\rfloor = N, \;\; 1 \le k \le p \tag{2} $$

The procedure of coarse-graining process is shown in Figure 1.

The coarse graining process has two main limitations. First, this process is not symmetric. According to Figure 1, for instance in scale 3, we could rationally expect the measure to behave the same for $y_{k,3}$ and $y_{k,4}$, in comparison with $y_{k,2}$ and $y_{k,3}$. However, at scale 3, $y_{k,1}$, $y_{k,2}$ and $y_{k,3}$ are separated from $y_{k,4}$, $y_{k,5}$ and $y_{k,6}$. Second, when the coarse graining process is computed, the number of samples of the resulting coarse-grained sequence is $\left\lfloor \frac{C}{\beta} \right\rfloor = N$. When the scale factor β is high, for each channel, the number of time sample points in the coarse-grained sequence decreases. This may yield unstable or undefined entropy values [30, 31]. To tackle these shortcomings, we propose the refined composite technique for multi-channel time series extending the previous definition for univariate signals [28]. The first step of refined composite multivariate multiscale entropy-based approaches is generating β coarse-grained multivariate time series $z_{\alpha}^{(\beta)} = \{x_{\alpha,k,i}{}^{(\beta)}\} \;\; 1 \le \alpha \le \beta$ where

$$ {}^{\mu}x_{\alpha,k,i}^{(\beta)} = \frac{1}{\beta} \sum_{b=(i-1)\beta}^{i\beta+\alpha-1} y_{k,b} \;\; 1 \le i \le \left\lfloor \frac{C}{\beta} \right\rfloor = N, \;\; 1 \le k \le p, \;\; 1 \le \alpha \le \beta \tag{3} $$

and

$$ {}^{\sigma^2}x_{\alpha,k,i}^{(\beta)} = \frac{1}{\beta}\sum_{b=(i-1)\beta+\alpha}^{i\beta+\alpha-1}(y_{k,b} - {}^{\mu}x_{\alpha,k,i}^{(\beta)})^2 \quad 1 \le i \le \left\lfloor \frac{C}{\beta} \right\rfloor = N,\ 1 \le k \le p,\ 1 \le \alpha \le \beta \tag{4} $$

were respectively used for RCmvMFE$_\mu$ and RCmvMFE$_\sigma{}^2$. As can be seen in Figure 2, in RCmvMFE$_\mu$/RCmvMFE$_\sigma{}^2$, for each scale factor β, we have β different multivariate signals $\mathbf{Z}_\alpha^{(\beta)}$, while in the mvMFE$_\mu$/mvMFE$_\sigma{}^2$ methods, only $\mathbf{Z}_1^{(\beta)}$ is considered. The second step of multivariate multiscale techniques is calculating multivariate fuzzy/sample entropy for each scale factor.

2.2. Multivariate Fuzzy Entropy

One of the most important shortcomings of the mvSE is that the method ignores every distance between two composite delay vectors (d) that is larger than a defined threshold r [15]. To alleviate this problem, a fuzzy membership function $\theta(d,r)$ [15] was proposed as follows:

$$ \theta(d,r) = \begin{cases} 1\ , & d \le r \\ \mathrm{e}^{-\ln(2)\left(\frac{d-r}{r}\right)^2}, & d > r \end{cases} \tag{5} $$

Although the above-mentioned problem is solved by using this function [15], this method is noticeably slower than the mvSE, especially when the number of channels or sample points of every channel, or the value of embedding dimension m is high. To tackle this deficiency, we propose to use another well-known fuzzy membership function as:

$$ \theta(d,r) = \exp\left(-(d)^{f_p} / r\right) \tag{6} $$

where f_p shows the fuzzy power and is usually equal to 2.

To calculate mvFE, multivariate embedded vectors are initially generated based on the Takens embedding theorem [15, 21, 32]. The multivariate embedded reconstruction is defined as:

$$X_m(i)=[x_{1,i},x_{1,i+\tau_1},...,x_{1,i+(m_1-1)\tau_1},x_{2,i},x_{2,i+\tau_2},...,x_{2,i+(m_2-1)\tau_2},...x_{P,i},x_{P,i+\tau_P},...,x_{P,i+(m_P-1)\tau_P}] \quad (7)$$

where $M=[m_1,m_2,...,m_p]$ and $\boldsymbol{\tau}=[\tau_1,\tau_2,...,\tau_p]$ are the embedding and the time lag vectors, respectively [29].

For *p*-variate time series $\{\mathbf{x}_k\}_{k=1}^{p}$, the mvFE algorithm, as a natural extension of the standard univariate fuzzy entropy, includes the following steps:

1. For each scale factor $1\le\alpha\le\beta$, form multivariate embedded vectors $X_m(i)\in R^m$ where i=1,2,...,N-n and n=max{M}×max{$\boldsymbol{\tau}$}.
2. Calculate the distance between any two composite delay vectors $X_m(i)$ and $X_m(j)$ as the maximum norm.
3. For a given threshold r and fuzzy power f_p, define a global quantity $\phi^m(r)$, as the average membership grade as:

$$\phi^m(r)=\frac{1}{(N-n)}\sum_{i=1}^{N-n}\frac{\sum_{j=1,i\neq j}^{N-n}\exp\left(\frac{-\left(d[X_m(i),X_m(j)]\right)^{f_p}}{r}\right)}{N-n-1} \quad (8)$$

4. Extend the dimensionality of the multivariate delay vector in (8) from m to (m+1). This can be done in p different ways, as from $[m_1,m_2,...,m_h,...,m_p]$ to $[m_1,m_2,...,m_{h+1},...,m_p]$ $(h=1,...,p)$. In the process, the dimension of the other variables are unchanged.
5. Calculate $\phi^{(m+1)}(r)$ which denotes the average over all $\phi^{(m_h+1)}(r)$ values in an (m+1)-dimensional space.
6. Finally, mvFE is defined as:

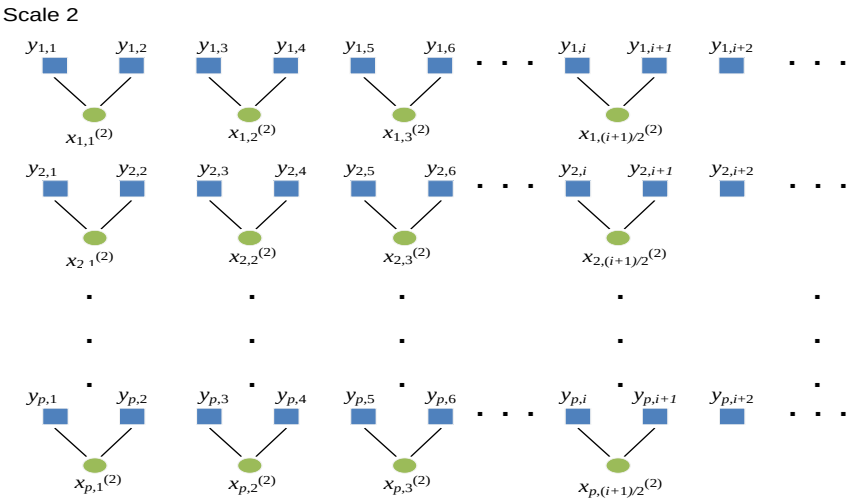
$$\mathrm{mvFE}(\mathbf{X},\boldsymbol{\tau},r,n)=-\ln\left(\frac{\phi^{(m+1)}(r)}{\phi^{m}(r)}\right) \tag{9}$$

3.3. Refined Composite Multivariate Fuzzy Entropy

As mentioned earlier, based on the proposed refined composite technique, for each scale factor β, we have β different multivariate signals $\mathbf{Z}_{\alpha}^{(\beta)}$. For each of $\mathbf{Z}_{\alpha}^{(\beta)}$, ${\phi_{\beta,\alpha}}^{m}|(\alpha=1,...,\beta)$ and ${\phi_{\beta,\alpha}}^{m+1}|(\alpha=1,...,\beta)$ are separately calculated. Next, the average of values of ${\bar{\phi}_{\beta,\alpha}}^{m}$ and ${\bar{\phi}_{\beta,\alpha}}^{m+1}$ on $1\leq\alpha\leq\beta$ are computed. Finally, the RCmvMFE is computed as follows:

$$\mathrm{RCmvMFE}(\mathbf{Y},\beta,M,n,r)=-\ln\left({\bar{\phi}_{\beta,\alpha}}^{m+1}\Big/{\bar{\phi}_{\beta,\alpha}}^{m}\right) \tag{10}$$

It should be added that because multi-channel signals may have different amplitude ranges, the distances calculated on embedded vectors may be biased toward the largest amplitude ranges variates. For this reason, we scale all the data channels to the same amplitude range and we normalize each data channel to unit standard deviation so that the total variation becomes equal to the number of channels or variables [21].



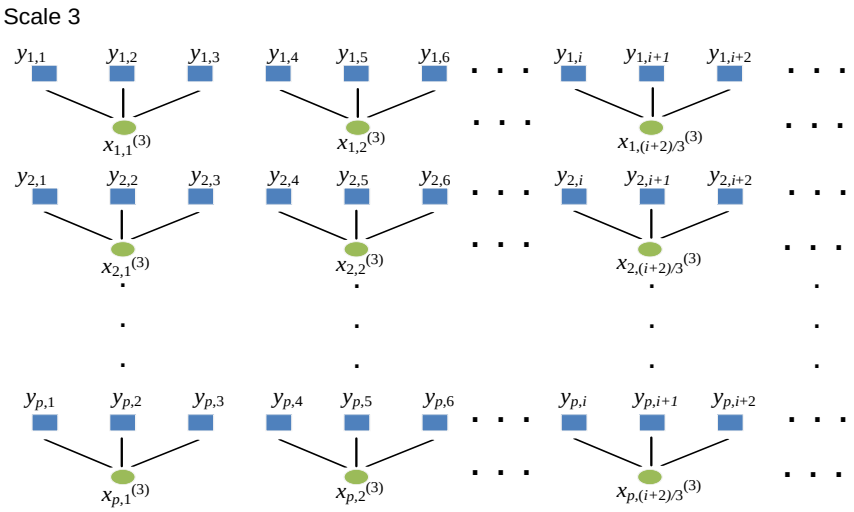
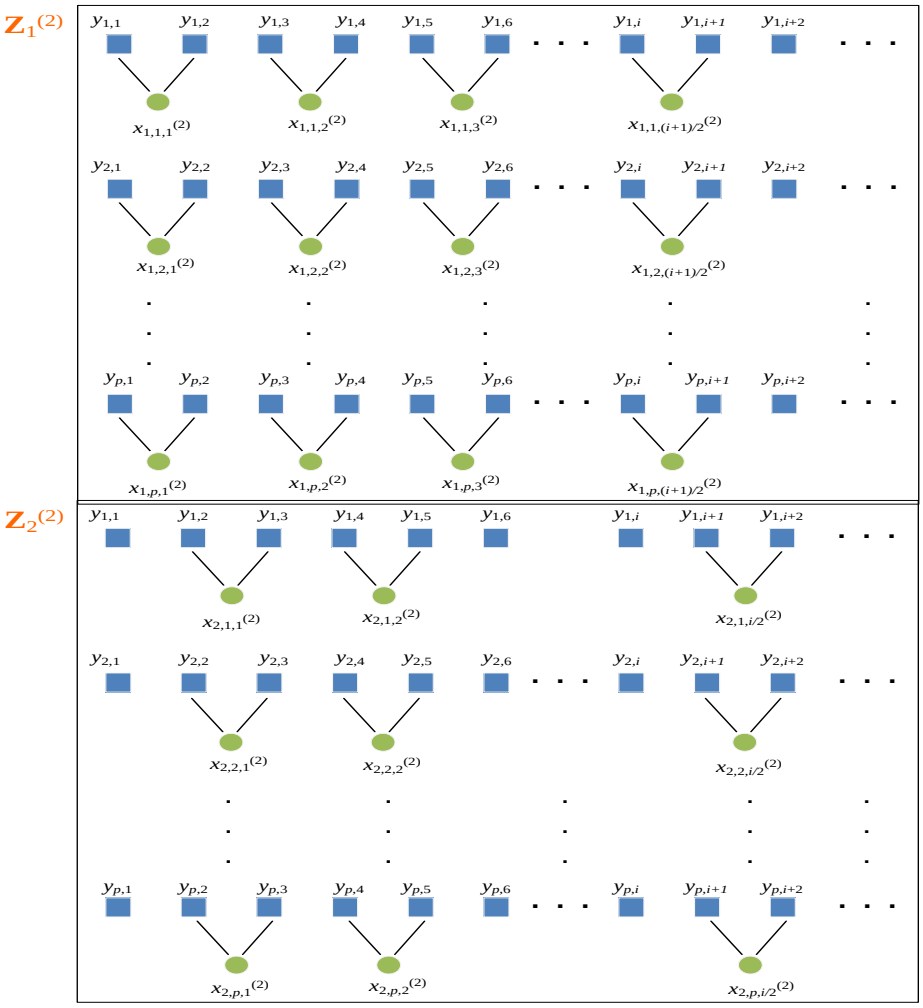


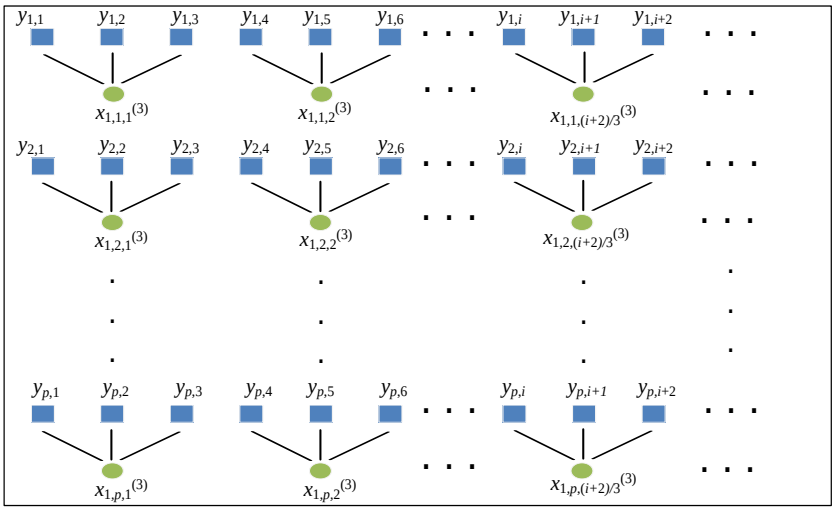
Figure 1. Illustration of the existing coarse-graining process for multivariate time series for scale factors 2 and 3. For each scale factor, one time series is computed.

Scale 2

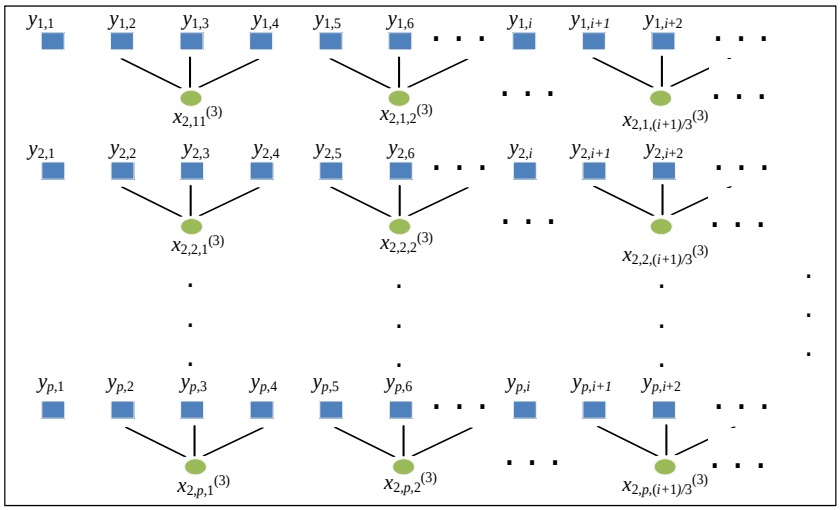

$\mathbf{Z}_1^{(2)}$
$y_{1,1}$ $y_{1,2}$ $y_{1,3}$ $y_{1,4}$ $y_{1,5}$ $y_{1,6}$ $y_{1,i}$ $y_{1,i+1}$ $y_{1,i+2}$
$x_{1,1,1}^{(2)}$ $x_{1,1,2}^{(2)}$ $x_{1,1,3}^{(2)}$ $x_{1,1,(i+1)/2}^{(2)}$
$y_{2,1}$ $y_{2,2}$ $y_{2,3}$ $y_{2,4}$ $y_{2,5}$ $y_{2,6}$ $y_{2,i}$ $y_{2,i+1}$ $y_{2,i+2}$
$x_{1,2,1}^{(2)}$ $x_{1,2,2}^{(2)}$ $x_{1,2,3}^{(2)}$ $x_{1,2,(i+1)/2}^{(2)}$
$y_{p,1}$ $y_{p,2}$ $y_{p,3}$ $y_{p,4}$ $y_{p,5}$ $y_{p,6}$ $y_{p,i}$ $y_{p,i+1}$ $y_{p,i+2}$
$x_{1,p,1}^{(2)}$ $x_{1,p,2}^{(2)}$ $x_{1,p,3}^{(2)}$ $x_{1,p,(i+1)/2}^{(2)}$
$\mathbf{Z}_2^{(2)}$
$y_{1,1}$ $y_{1,2}$ $y_{1,3}$ $y_{1,4}$ $y_{1,5}$ $y_{1,6}$ $y_{1,i}$ $y_{1,i+1}$ $y_{1,i+2}$
$x_{2,1,1}^{(2)}$ $x_{2,1,2}^{(2)}$ $x_{2,1,i/2}^{(2)}$
$y_{2,1}$ $y_{2,2}$ $y_{2,3}$ $y_{2,4}$ $y_{2,5}$ $y_{2,6}$ $y_{2,i}$ $y_{2,i+1}$ $y_{2,i+2}$
$x_{2,2,1}^{(2)}$ $x_{2,2,2}^{(2)}$ $x_{2,2,i/2}^{(2)}$
$y_{p,1}$ $y_{p,2}$ $y_{p,3}$ $y_{p,4}$ $y_{p,5}$ $y_{p,6}$ $y_{p,i}$ $y_{p,i+1}$ $y_{p,i+2}$
$x_{2,p,1}^{(2)}$ $x_{2,p,2}^{(2)}$ $x_{2,p,i/2}^{(2)}$

Scale 3

$\mathbf{Z}_1^{(3)}$



$\mathbf{Z}_2^{(3)}$



$\mathbf{Z}_3^{(3)}$

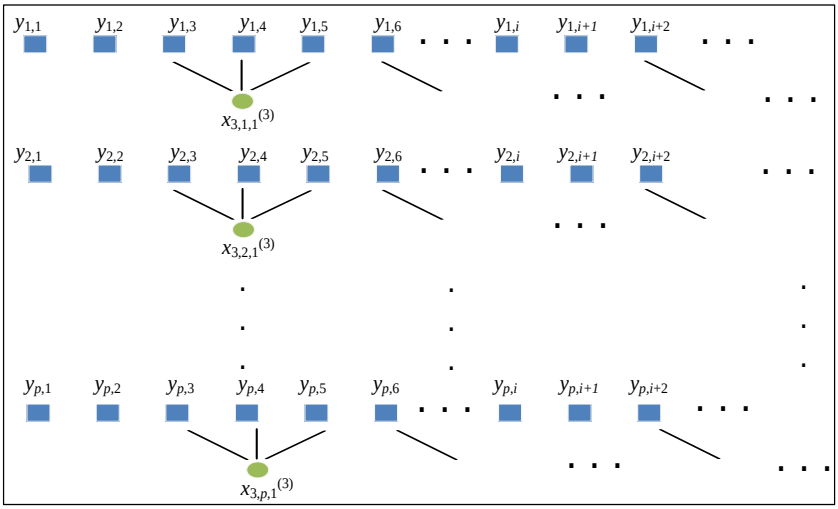


Figure 2. Illustration of the proposed coarse-graining process (refined composite technique) for multivariate time series for scale factors 2 and 3. For scale factor β, there are β different time series.

2.3. Multivariate Entropy Parameters

The time lag is an important parameter for (multivariate) multiscale entropy-based methods, since an appropriate choice of time lag can make the necessary embedding dimension lower [13, 21, 32]. There is not any proven criterion for choosing this parameter. Generally, we have a fixed time lag and then, adjust the embedding dimension accordingly [24]. There exist a few approaches to choose the time lag for a univariate signals, such as mutual information and autocorrelation [32-34]. For multivariate signals, we can also use these approaches to find $\boldsymbol{\tau}$ for each channel separately. Moreover, we can extend these methods for multivariate time series [32]. Nevertheless, for simplicity, we, like the proposers of the mvMSE_{μ} [21, 29], choose τ_k=1. The sensitivity of mvMSE_{μ} and the proposed method RCmvMFE_{μ} to the signal length will be described in Section 4. m_k, f_p and r for all of the approaches are respectively chosen 2, 2 and 0.15 multiplied by the standard deviation (SD) of the original time series according to [2, 21]. Our choice of the value of r is further supported by the fact that the behavior of the results obtained with a different value of r, as in [20], is similar to ours [21].

3. Evaluation Signals

White Gaussian (discrete-time) noise (WGN) and $1/f$ noise are two most important signals to evaluate the multiscale entropy-based approaches for univariate and multivariate signals [13, 21, 24].

3.1. Synthetic Signals

WGN is a random signal which has equal energy at all frequencies. The name white originates from the fact that this kind of signal has a constant power spectral density $S(f)$ as follows:

$$S(f) = C_w \tag{11}$$

where C_w is a constant number [35]. WGN signal is defined as a sequence of consecutively uncorrelated random variables with zero mean and finite variance [36].

A stochastic process that is appropriate to model evolutionary or developmental systems, characterized by equal energy per octave, is called pink noise with power spectral density as follows:

$$S(f) = \frac{C_f}{f} \tag{12}$$

where C_f is a constant. As can be seen in Equation 12, the power spectrum density of pink noise is inversely proportional to frequency [35].

To generate the correlated 2-channel noise signals, first we created a two-channel uncorrelated random signals **H**. Multiplying **H** with the standard deviation (hereafter, sigma) and adding the value of the mean (hereafter, mu) was the next step. Then, we multiplied **H** with the upper triangular matrix **L** obtained from the Cholesky decomposition of a defined correlation matrix **R** (which is symmetric and positive) to set the correlation. A comprehensive study on the effect of correlated and uncorrelated WGN and 1/*f* noise on multiscale entropy-based methods was done in [13, 37].

3.2. Real Signals

The proposed methods RCmvMFEσ^2 and RCmvMFE$_\mu$ are next assessed by using two publicly available datasets: bivariate intracranial focal and non-focal EEG time series and trivariate Fantasia signals from young and elderly subjects.

3.2.1. Intracranial EEG Data

The intracranial EEG signals were recorded from five patients suffering from pharmacoresistant focal-onset epilepsy leading to two main separate sets of signals. The first one was recorded from brain regions where the primarily ictal EEG recordings changes were detected as judged by expert visual inspection ("focal signals"). The second set of signals was recorded from brain regions not involved at seizure onset ("non-focal signals"). The focal and non-focal EEG time series are recorded from patients affected by focal epilepsy. Focal epilepsy influences on only a limited part of the patient's brain. The focal EEG recording are acquired outside of seizures from the channels that first detected ictal EEG time series changes by visual inspection. The remaining EEGs are termed as non-focal EEGs [38, 39]. For each patient, there are 750 focal and 750 non-focal bivariate time series. The length of each signal was 10240 sample points: 20 seconds at a sampling frequency of 512 Hz. Each pair includes two EEG time series recorded from adjacent channels. For more information about the dataset, please, refer to [38]. Before computing the existing and proposed approaches, all signals were digitally filtered employing an FIR band-pass filter with cut-off frequencies at 0.5 Hz and 40 Hz.

3.2.2. Fantasia Database

A subset of the Fantasia database was chosen which includes 10 young (21 - 34 years old) and 10 old (68 - 85 years old), rigorously-screened healthy individuals underwent 120 minutes of

continuous supine resting while continuous electrocardiographic (ECG), respiration signals, and uncalibrated continuous non-invasive blood pressure recordings, were collected. Each group consisted of 5 women and 5 men [40]. All 20 individuals remained in an inactive state in sinus rhythm when watching the movie Fantasia (Disney, 1940) to help to maintain wakefulness. For each subject, all three time series were digitized at 250 Hz [40].

4. Results and Discussions

4.1. Synthetic Signals

We first apply the proposed and conventional approaches to white and $1/f$ noise as two kinds of signals that are used to evaluate the performance of multivariate/univariate multiscale entropy-based methods [21, 29, 41]. The results are shown in Figures 3 to 5. The number of sample points of each of the WGN and $1/f$ noise was 20000.

Figures 3(a) and (b) illustrate the results obtained for mvMSE_{μ} using 40 different 2- and 5-channel WGN and $1/f$ noise signals, respectively. The results are consistent with the fact that $1/f$ noise has more complex structure than WGN [13, 17, 21]. For 2-channel noise time series and for scales 1 and 2, the entropy values of WGN signals are higher than those of $1/f$ noise. However, for $2 < \beta$, when the entropy value for the coarse-grained $1/f$ noise signal stays almost constant, that for the coarse-grained WGN signal monotonically decreases. For WGN, when the length of the signal, obtained by the coarse-graining process, decreases (i.e., the scale factor increases), the mean value of inside each signal converges to a constant value and the SD becomes smaller. Therefore, no new structures are revealed on higher scales. This demonstrates WGN time series has information only in small time scales [13]. In contrast, for $1/f$ noise time series, the mean value of the oscillations inside each signal does not converge to a constant value

[13, 14]. Figure 3(b) shows that for 5-channel noise time series and for each of $1<\beta$, the mean of entropy values for $1/f$ noise is larger than that for WGN. This fact may be considered as a relative advantage of the larger number of channels.

To demonstrate the importance of refined composite technique for $mvMSE_{\mu}$, the results obtained by $RCmvMSE_{\mu}$ algorithm for 40 different 2- and 5-channel noise time series are shown in Figures 3(c) and (d), respectively. To compare the results obtained by the multivariate approaches, we used the coefficient of variation (CV) defined as the SD divided by the mean [42]. The main reason to use this measure is that the SDs of some signals generally increase or decrease proportionally as the mean increases or decreases. Therefore, the CV, as a standardization of the SD, allows comparison of variability estimates regardless of the magnitude of analyte concentration [42]. We consider only the $1/f$ noise time series results at the highest scale factor, i.e. 60, because the average values of WGN results at this scale are near 0 and it may lead to unreliable CV values. As can be seen in Figures 3(a) to (d), the results obtained by $RCmvMSE_{\mu}$ are more stable than those obtained by $mvMSE_{\mu}$. This fact is also confirmed by the smaller CV values for $RCmvMSE_{\mu}$, shown in Table 1.

To depict the importance of FuzEn in multivariate multiscale methods, the behavior of $mvMFE_{\mu}$ is considered using WGN and $1/f$ noise time series. Figures 3(e) and (f) demonstrate, respectively, the results obtained by 2- and 5-channel signals using $mvMFE_{\mu}$. Comparing Figures 3(a) and (b) with Figures 3(e) and (f) shows that the FuzEn improves the stability of $mvMSE_{\mu}$. $RCmvMFE_{\mu}$ results using 2- and 5-channel WGN and $1/f$ noise time series are shown in Figures 3(g) and (h), in that order. The results illustrate that $RCmvMFE_{\mu}$, like $RCmvMSE_{\mu}$, can improve the stability of its basic method ($mvMFE_{\mu}$). It is worth mentioning that although the results obtained by the fuzzy membership function used in this paper is similar to those achieved by the

fuzzy membership function proposed in [15], the running time for the existing $mvMFE_{\mu}$ is about six times more than that for the proposed one. As expected, the CV values of the $mvMFE_{\mu}$ results are noticeably smaller than those of $mvMSE_{\mu}$ ones. Comparing $RCmvMFE_{\mu}$- and $mvMFE_{\mu}$-based CV values demonstrates the importance of refined composite technique in multi-channel signals.

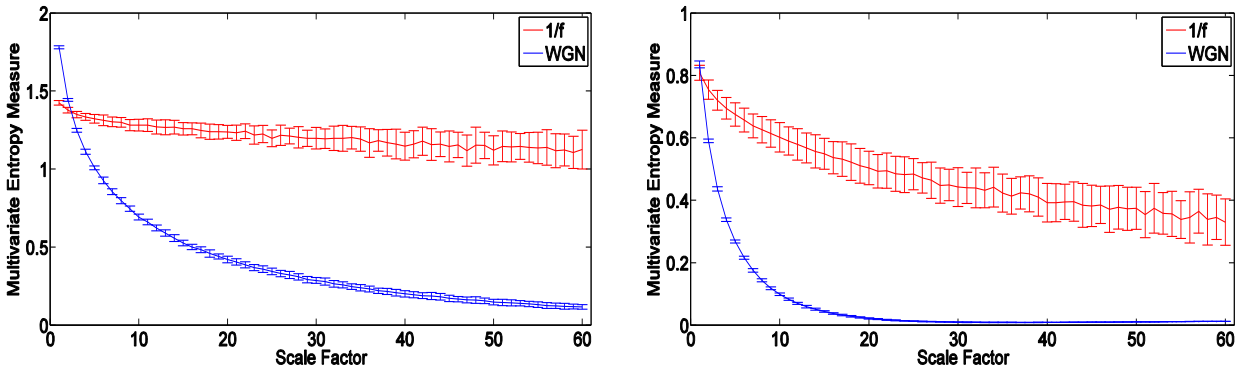
The results obtained by the $mvMPE_{\mu}$ method for 2- and 5-channel 1/*f* noise and WGN signals are shown in Figures 4(a) and (b), respectively. The results demonstrates the $mvMPE_{\mu}$ method, though powerful and widely-used, do not follow the concept of complexity about 1/*f* and WGN [13, 17, 18]. Note that, we chose m=5 for the $mvMPE_{\mu}$ according to [43].

As a recent interest in using other statistical moments, we employ variance, instead of mean, in the coarse-graining process. The results are shown in Figures 5(a) and (b) using $mvMSE_{\sigma^2}$ for respectively 2- and 5-channel WGN and 1/*f* noise time series. Although the crossing points of the curves of WGN and 1/*f* noise results appear in higher scale factors in comparison with $mvMSE_{\mu}$, the general behaviors of curves obtained by $mvMSE_{\mu}$ and $mvMSE_{\sigma^2}$ are similar. In addition, the results for 5-channel time series have smaller SD and the profiles for both conditions crossed at lower scale, compared with 2-channel time series. This again demonstrates the importance of using multivariate signals. As mentioned before, the signal **X** is more complex than the signal **Y**, if for most time scales, the mvSE/mvFE values for time series **X** are higher than those for time series **Y** [21, 24]. Thus, the number of scales at which the multivariate entropy value of **X** is higher than that of **Y** is relevant. In this particular case, WGN is less irregular than 1/*f* noise for 2- and 5-channel noise signals from scale factor 59 and 44, respectively. Thus, a smaller number of temporal scales are needed to reveal the multiscale complexity of the signal for the 5-channel noise signals in comparison with the 2-chanel noise. Thus, this fact and the smaller SD values

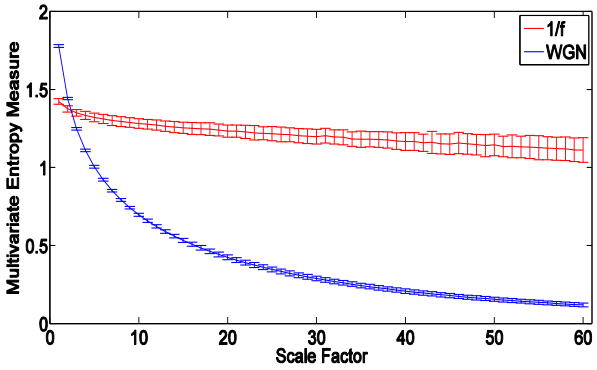
obtained for the 5-channel signal support the use of a larger number of channels for variance-based coarse-graining.

It is worth mentioning that for each scale factor, the SD of results using $mvMSE_{\mu}$ is noticeably larger than $mvMSE_{\sigma^2}$. Figures 5(c) and (d) show the results obtained by $RCmvMSE_{\sigma^2}$ for 2- and 5-channel signals, respectively. As can be seen in Table 1, the CV values of the $RCmvMSE_{\sigma^2}$ results are smaller than those of $mvMSE_{\sigma^2}$ ones, something that shows the importance of refined composite algorithm to increase the stability of the results.

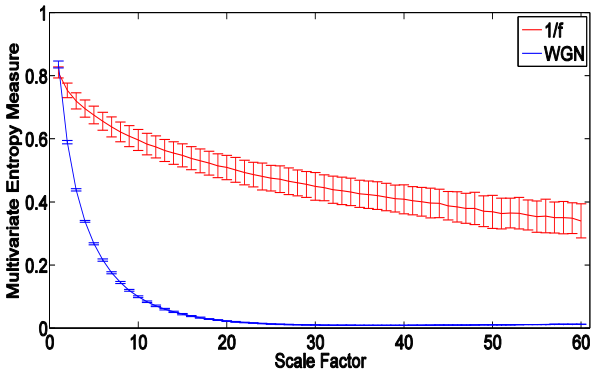
Figures 5(e) and (f) respectively show the results obtained by 2- and 5-channel signals based on $mvMFE_{\sigma^2}$. The general trends of these results are the same as Figures 5(a) and (b), although fuzEn-based results are noticeably more stable than SampEn-based ones. This is further supported by the CV vales depicted in Table 1. In addition, the crossing points between the WGN and 1/*f* noise results using mvMFE methods are placed in smaller scale factors in comparison with those using $mvMSE_{\sigma^2}$. To improve the stability of the $mvMFE_{\sigma^2}$ method, we propose to use refined composite technique. The results are shown in Figures 5(g) and (h). As can be seen in these figures, for each time scale, the CV of $RCmvMFE_{\sigma^2}$ values is smaller than that of $mvMFE_{\sigma^2}$ measures. It is worth noting that the advantages of mvMFE over the exiting method mvMSE are still into account for RCmvMFE.



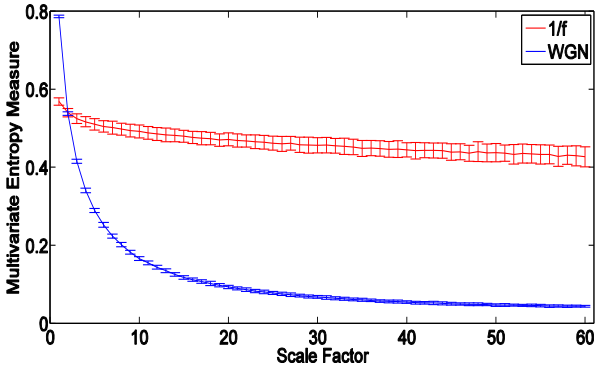
(a) $mvMSE_{\mu}$ (number of channels=2)



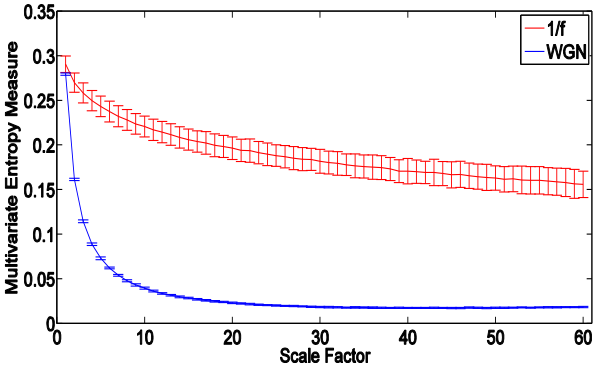
(b) $mvMSE_{\mu}$ (number of channels=5)



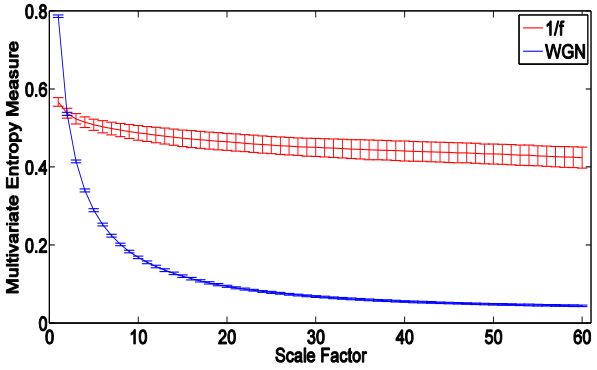
(c) $RCmvMSE_{\mu}$ (number of channels=2)



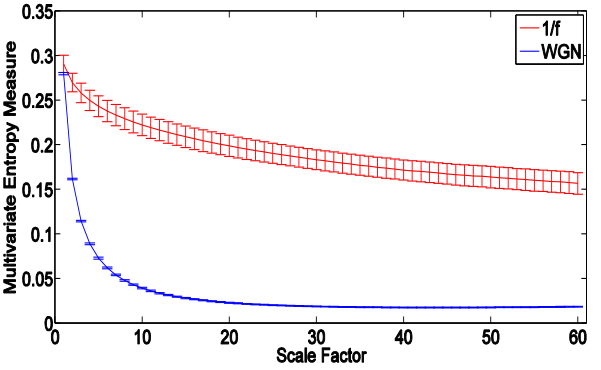
(d) $RCmvMSE_{\mu}$ (number of channels=5)



(e) $mvMFE_{\mu}$ (number of channels=2)



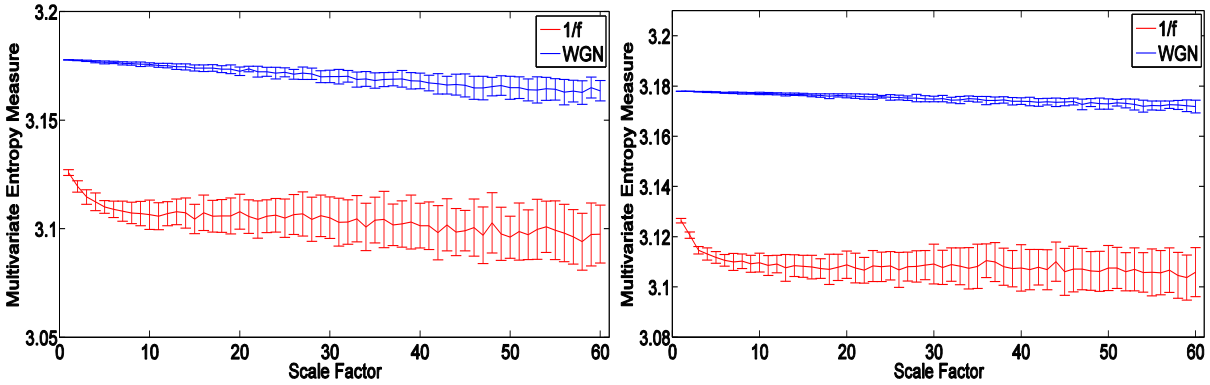
(f) $mvMFE_{\mu}$ (number of channels=5)



(g) $RCmvMFE_{\mu}$ (number of channels=2)

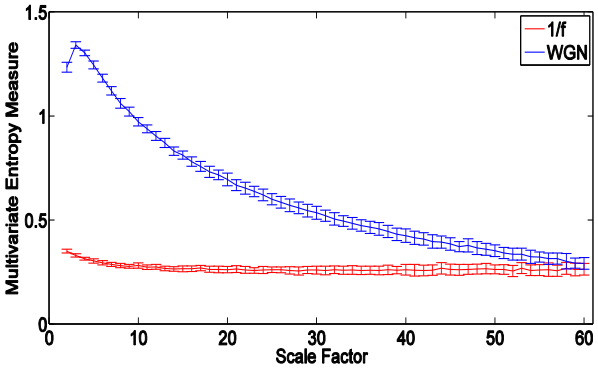
(h) $RCmvMFE_{\mu}$ (number of channels=5)

Figure 3. Mean value and SD of results of the (a) $mvMSE_{\mu}$ (number of channels=2), (b) $mvMSE_{\mu}$ (number of channels=5), (c) $RCmvMSE_{\mu}$ (number of channels=2), (d) $RCmvMSE_{\mu}$ (number of channels=5), (e) $mvMFE_{\mu}$ (number of channels=2), (f) $mvMFE_{\mu}$ (number of channels=5), (g) $RCmvMFE_{\mu}$ (number of channels=2), and (h) $RCmvMFE_{\mu}$ (number of channels=5) computed from 40 different multichannel WGN and 1/*f* noise times series. Red and blue indicate 1/*f* and WGN results, respectively.

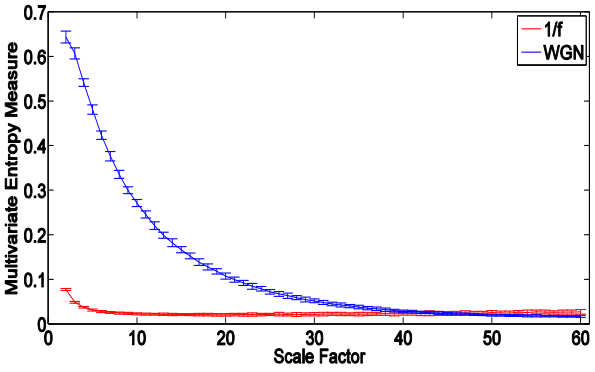


(a) $mvMPE_{\mu}$ (number of channels=2) *(b) $mvMPE_{\mu}$ (number of channels=5)*

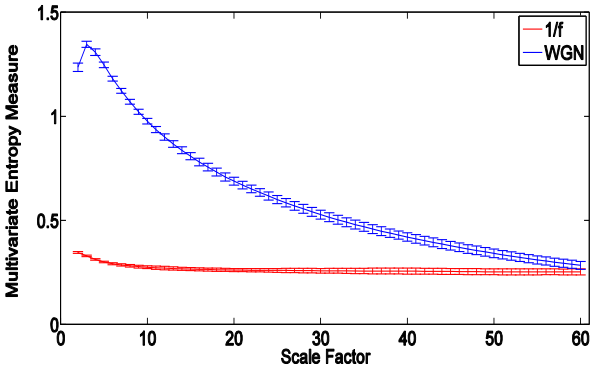
Figure 4. Mean value and SD of results of the (a) $mvMPE_{\mu}$ (number of channels=2), and (b) $mvMPE_{\mu}$ (number of channels=5) computed from 40 different multichannel WGN and 1/f noise times series. Red and blue indicate 1/f and WGN results, respectively.



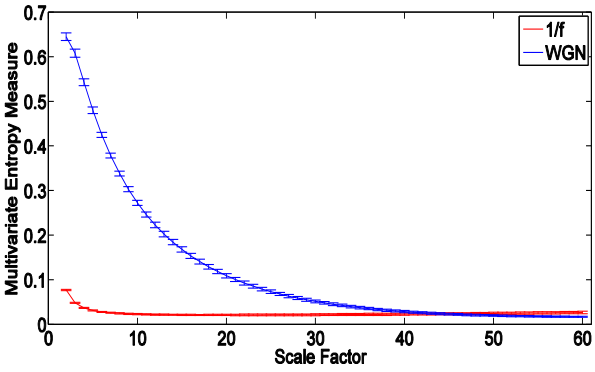
(a) $mvMSE_{\sigma}^{2}$ (number of channels=2)



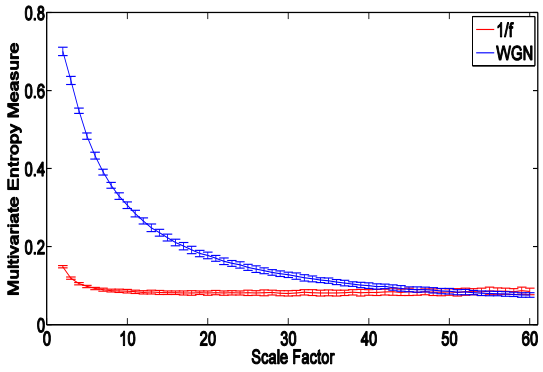
(b) $mvMSE_{\sigma}^{2}$ (number of channels=5)



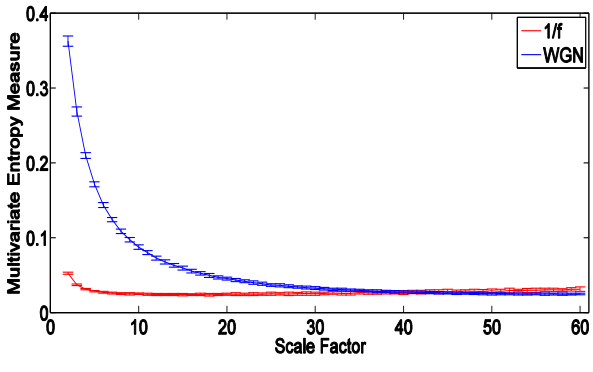
(c) $RCmvMSE_{\sigma}^{2}$ (number of channels=2)



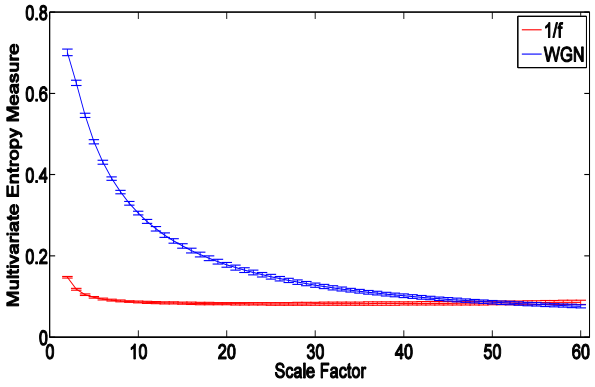
(d) $RCmvMSE_{\sigma}^{2}$ (number of channels=5)



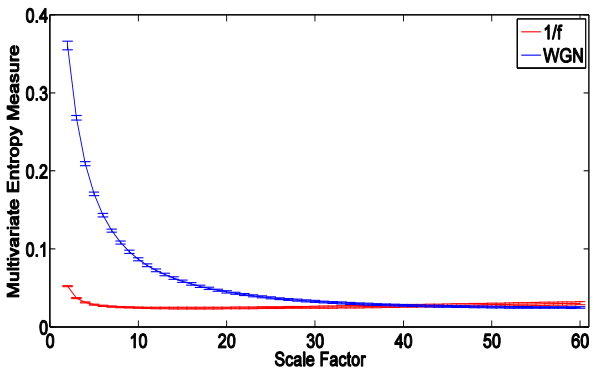
(e) $mvMFE_{\sigma}^{2}$ (number of channels=2)



(f) $mvMFE_{\sigma}^{2}$ (number of channels=5)



(g) RCmvMFE$_{\sigma}^2$ (number of channels=2)



(h) RCmvMFE$_{\sigma}^2$ (number of channels=5)

Figure 5. Mean value and SD of results of the (a) mvMSE$_{\sigma}^2$ (number of channels=2), (b) mvMSE$_{\sigma}^2$ (number of channels=5), (c) RCmvMSE$_{\sigma}^2$ (number of channels=2), (d) RCmvMSE$_{\sigma}^2$ (number of channels=5), (e) mvMFE$_{\sigma}^2$ (number of channels=2), (f) mvMFE$_{\sigma}^2$, (number of channels=5), (g) RCmvMFE$_{\sigma}^2$ (number of channels=2), (h) RCmvMFE$_{\sigma}^2$, (number of channels=5) computed from 40 different multichannel WGN and 1/*f* noise times series. Red and blue indicate 1/*f* and WGN results, respectively.

Table 1: The CV values of the proposed and existing multivariate multiscale entropy-based results to compare the normalized SD values at the scale factor 60.

Methods	2-channel time series	5-channel time series
mvMSE$_{\mu}$	0.1116	0.2272
RCmvMSE$_{\mu}$	0.0719	0.1612
mvMFE$_{\mu}$	0.0584	0.0961

$RCmvMFE_{\mu}$	0.0392	0.0778
$mvMSE_{\sigma^2}$	0.1826	0.1907
$RCmvMSE_{\sigma^2}$	0.0575	0.1019
$mvMFE_{\sigma^2}$	0.939	0.0966
$RCmvMFE_{\sigma^2}$	0.0522	0.0480

Table 2 demonstrates that the running time values of the proposed and conventional methods for 2- and 5-cahnnel signals with the length 1000, 3000, and 10000 sample points. In this paper, the simulations have been carried out using a PC with Intel (R) Xeon (R) CPU, E5420, 2.5 GHz and 8-GB RAM by MATLAB R2010a. $mvMSE_{\mu}$ is about 3 times faster than the $RCmvMSE_{\mu}$ for either 2- or 5-channel time series with different sample points. The running time of the existing $mvMFE_{\mu}$ [15] is around three and six times more than that of the proposed $mvMFE_{\mu}$ with C=1000 and C=10000, respectively. This shows the far superiority of the $mvMFE_{\mu}$ in terms of calculation time. The proposed $mvMFE_{\mu}$ is comparatively as fast as $mvMSE_{\mu}$ and therefore, the running time of $RCmvMFE_{\mu}$ is near to that of $RCmvMSE_{\mu}$. Since the multivariate multiscale entropy algorithms based on variance starts from scale factor 2, the running time of this kind of methods is noticeably smaller than that of the multivariate multiscale entropy algorithms that use mean in their coarse-graining process. For variance-based multivariate multiscale algorithms, like the average-based ones, the running time values of the $mvMSE_{\sigma^2}$ and $mvMFE_{\sigma^2}$ are relatively equal.

Table 2: The computation time of the multivariate multiscale entropy-based proposed and existing methods with β_{max}=10.

Methods	1000 sample points (2 channels)	1000 sample points (5 channels)	3000 sample points (2 channels)	3000 sample points (5 channels)	10000 sample points (2 channels)	10000 sample points (5 channels)
$mvMSE_{\mu}$	0.271 s	1.146 s	1.274 s	7.051 s	8.281 s	54.108 s
$RCmvMSE_{\mu}$	0.982 s	3.316 s	3.807 s	16.71 s	22.18 s	124.251 s
$mvMFE_{\mu}$ [15]	0.753 s	3.846 s	4.459 s	28.28 s	50.65 s	269.160 s
Proposed $mvMFE_{\mu}$	0.278 s	1.148 s	1.238 s	6.812 s	7.612 s	49.214 s
$RCmvMFE_{\mu}$	1.001 s	3.329 s	3.796 s	16.21 s	18.147 s	105.16 s
$mvMSE_{\sigma^2}$	0.071 s	0.393 s	0.391 s	2.34 s	3.32 s	19.41 s
$RCmvMSE_{\sigma^2}$	0.306 s	1.564 s	1.481 s	9.512 s	12.301 s	67.421 s
$mvMFE_{\sigma^2}$	0.071 s	0.352 s	0.362 s	2.173 s	2.881 s	18.473 s
$RCmvMFE_{\sigma^2}$	0.310 s	1.546 s	1.451 s	8.691 s	9.26 s	60.365 s

To evaluate the sensitivity of $mvMSE_{\mu}$ and the proposed method $RCmvMFE_{\mu}$ to the signal length, we consider 3-channel WGN and 1/*f* noise signals as functions of sample points size *C*, where *m*=2. Figures 6(a) to (F) show the $mvMSE_{\mu}$ values for the signal length 100, 300, 1000, 3000, 10000, and 30000, respectively, computed from 40 different multichannel WGN and 1/*f* noise times series. The corresponding $RCmvMFE_{\mu}$ results are depicted in Figure 7. The results

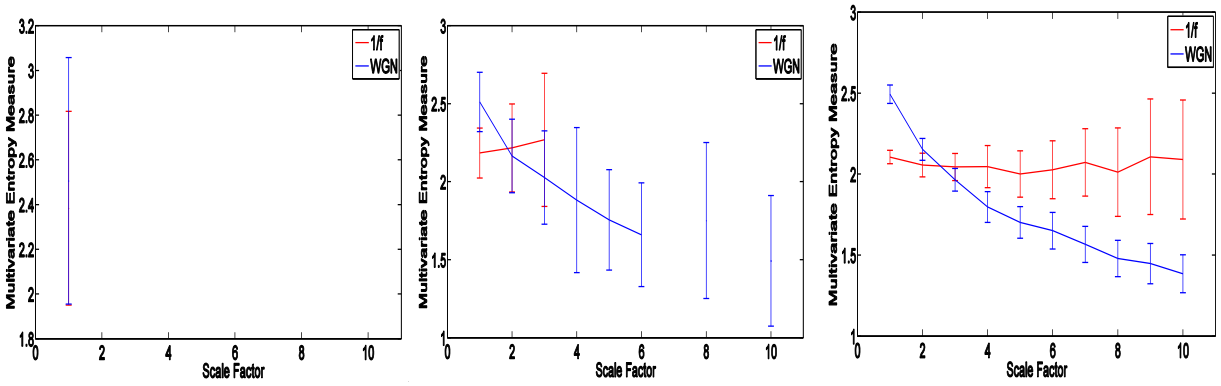
show the greater the value of C, the more robust the both $mvMSE_{\mu}$ and $RCmvMFE_{\mu}$ estimates, as seen from the error bars. As expected theoretically [13], for each C, $1/f$ noise is more complex than WGN using both of the methods.

It has been recommended that the number of sample points is at least 10^m, or preferably at least 30^m, to robustly estimate ApEn or SampEn in univariate signals [44]. Because the coarse-graining step decreases the signal length by the scale factor β, and here we have β_{max}=10 and m=2, the original signal should have at least 1000 sample points. Moreover, in mvMSE, the number of instances where $d = [X_m(i), X_m(j)] \leq r,\ j \neq i$ is counted. In case the signal length is too small, this number may be 0, leading to an undefined entropy value. Based on this fact, the results obtained by $mvMSE_{\mu}$ for C=100 and 300, respectively shown in Figures 6(a) and (b), are not defined. In contrast, the fuzzy entropy-based methods do not count matches, but consider all possible range of distances between any two composite delay vectors $X_m(i)$ and $X_m(j)$. Thus, mvMFE avoids resulting in undefined entropy values in such situations.

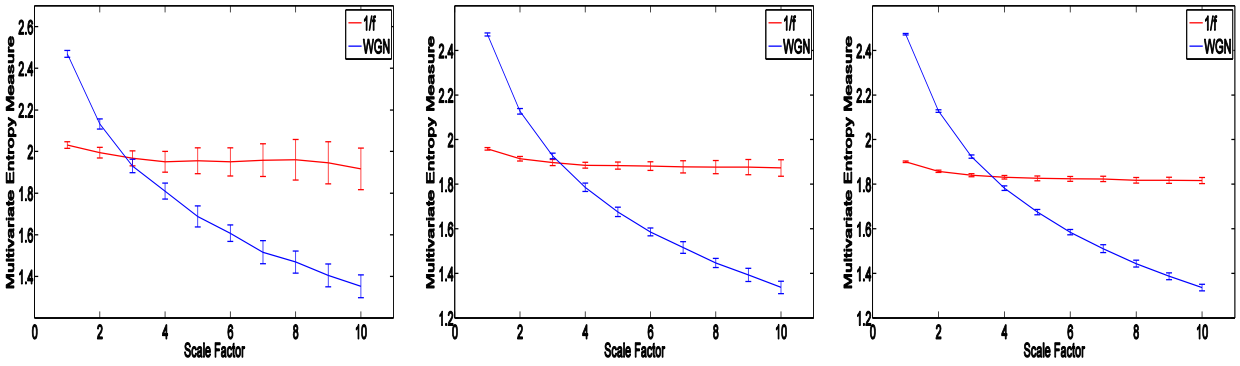
We now consider the case of $RCmvMFE_{\mu}$/$RCmvMSE_{\mu}$ at scale factor β. Although the length of the signal decreases β times, we now consider β time coarse-grained time series, instead of only one time series as in existing multivariate multiscale entropy-based methods. Thus, in refined composite-based methods, we have β times more comparison than the number of instances in $mvMFE_{\mu}$/$mvMSE_{\mu}$, leading to more reliable results, especially for short signals. In fact, in the refined composite-based techniques, the effect of the shorter coarse grained sequences would tend to cancel out. All in all, using the fuzzy membership function and/or refined composite technique cause the $RCmvMFE_{\mu}$ to be more reliable for short signals. The results obtained by the $RCmvMFE_{\mu}$ have considerably smaller SD than those obtained by $mvMSE_{\mu}$. This shows the superiority of $RCmvMFE_{\mu}$ over $mvMSE_{\mu}$. It is worth noting that the results obtained by

$mvMSE_{\sigma}^{2}$ and $RCmvMFE_{\sigma}^{2}$ are similar to Figures 6 and 7, although like the abovementioned results obtained by multivariate entropy methods based on the variance, the crossing points are different.

The number of embedding vectors for multivariate- and univariate-based methods are equal. Given that in multivariate/univariate entropy methods, we consider the Euclidean difference between pairs of embedding vectors, the number of matches is similar to that obtained by its univariate methods. Based on this fact and the experimental study on different signals, Ahmed and Mandic recommended that multivariate signals should have 100 or preferably 300 sample points, regardless of their number of channels [21, 24, 29]. Note that, although the number of samples for each of these vectors for multivariate methods is *p* times larger than that for SampEn/FuzEn, this is not as much of a problem in Fuzzy approaches, but the phenomenon may still occur.

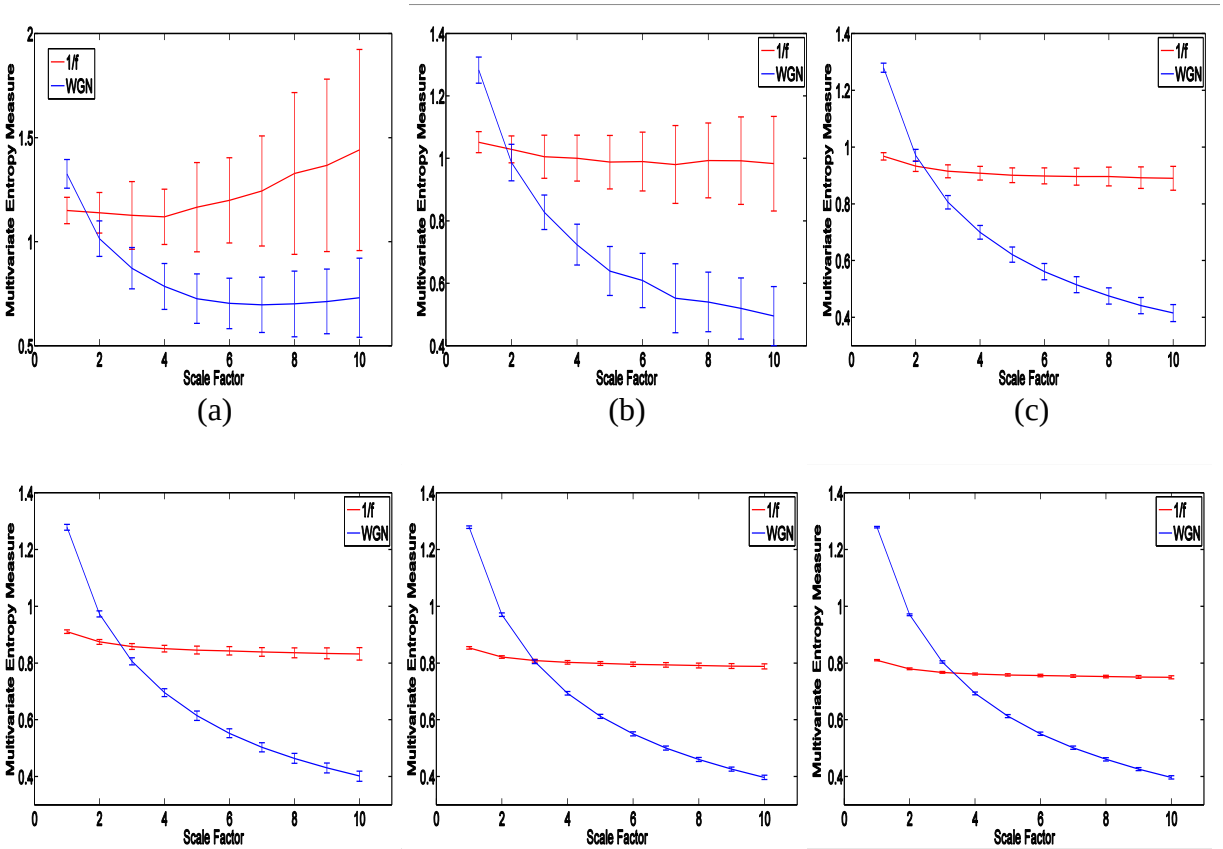


(a) (b) (c)



(d) (e) (f)

Figure 6. mvMSE$_\mu$ as a function of data length C, (a) C=100, (b) C=300, (c) C=1000, (d) C=3000, (e) C=10000, and (f) C=30000 computed from 40 different multichannel WGN and 1/f noise times series. Red and blue indicate 1/f and WGN results, respectively.



(a) (b) (c)

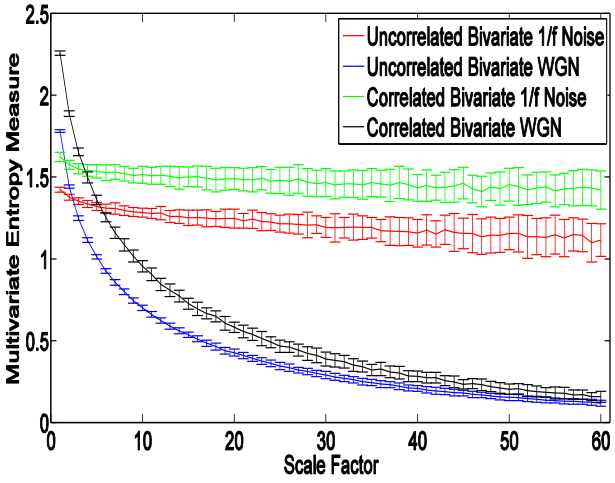
(d) (e) (f)

Figure 7. RCmvMFE$_\mu$ as a function of data length C, (a) C=100, (b) C=300, (c) C=1000, (d)

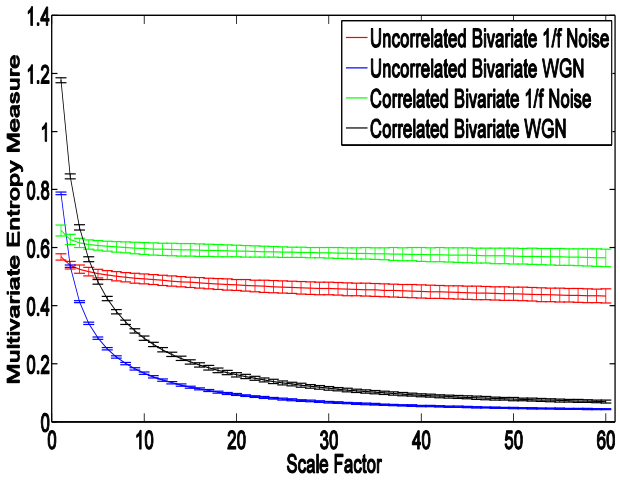
C=3000, (e) *C*=10000, and (f) *C*=30000 computed from 40 different multichannel WGN and 1/*f* noise times series. Red and blue indicate 1/*f* and WGN results, respectively.

Correlated noise has more structural complexity in comparison with uncorrelated noise and shows relatively higher regularity because the majority of entropy-based approaches fail to take into account the multiple temporal and spatial scales of the system [13, 37]. To this end, Costa *et al.* introduced multiscale entropy to measure the entropy values in different time scales [14]. As mentioned earlier, MSE- and MFE-based algorithms consider the long-term correlations within a univariate signal, but because of their univariate nature, they cannot model the cross-channel information existing in multivariate time series. In contrast, mvMSE and mvMFE are designed for multivariate time series. To show this fact, we created different combinations of bivariate WGN and 1/*f* noise signals (according to [21]), making the channels correlated. A brief algorithm to create uncorrelated noise signals was described in Section 3. Figures 8(a), (b), (c), and (d) respectively depict the ability of mvMSE_{μ}, RCmvMFE_{μ}, mvMSE_{σ^2}, and $\text{RCmvMFE}_{\sigma^2}$ to model both the within- and cross-channel properties in multivariate time series. As can be seen in Figure 8, the correlated bivariate 1/*f* noise is more complex at high scale factors using all four approaches, followed by the uncorrelated 1/*f* noise, and correlated and uncorrelated WGN, although the ordering of the entropy values are different in small scale factors. Thus, as desired, all mentioned multivariate multiscale entropy-based methods tackle both the cross- and within-channel correlations. This fact is also in agreement with the results reported in [29]. Here, we chose $\mathbf{R}=\begin{bmatrix} 1 & 0.95 \\ 0.95 & 1 \end{bmatrix}$, according to [21] . It is worth noting that, if we change **R** to make the signals less correlated (e.g., $\mathbf{R}=\begin{bmatrix} 1 & 0.65 \\ 0.65 & 1 \end{bmatrix}$), the entropy values decrease, as expected. We

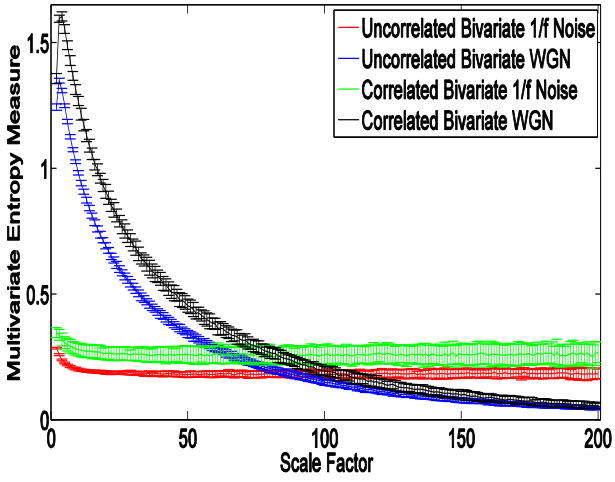
consider the $mvMSE_{\mu}$ and $RCmvMFE_{\mu}$ curves until the scale factor 60, because the ordering of curves do not change after this scale factor. Another important point about Figure 8 is demonstrating the importance of refined composite technique and fuzzy-based entropy to improve the stability of the results.



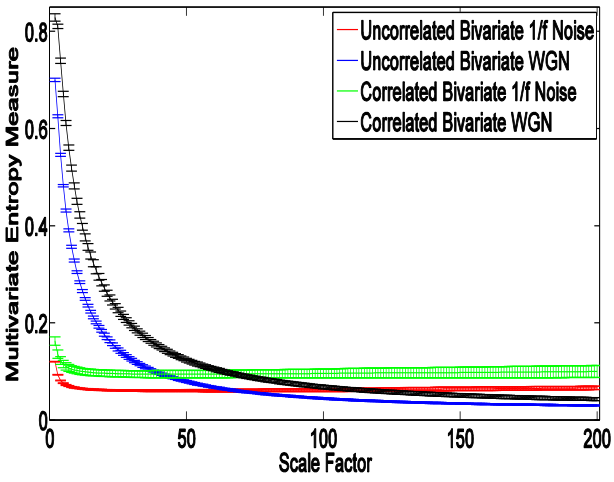
(a) $mvMSE_{\mu}$



(b) $RCmvMFE_{\mu}$



(c) $mvMSE_{\sigma}^{2}$



(d) $RCmvMFE_{\sigma}^{2}$

Figure 8. Mean value and SD of results of the (a) $mvMSE_{\mu}$, (b) $RCmvMFE_{\mu}$, (c) $mvMSE_{\sigma^2}$, and (d) $RCmvMFE_{\sigma^2}$ computed from 40 different bivariate WGN and 1/*f* time series, each has 40000 sample points.

4.2. Real Signals

We evaluate the suitability of the $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma^2}$ approaches to characterize the focal and non-focal EEG signals. We averaged the results over patients for each of two groups. The error bars demonstrating the SD and mean of the $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma^2}$ values computed from focal and non-focal bivariate EEG time series are depicted in Figures 9(a) and (b), respectively. The average entropy of non-focal EEG time series is larger than that of focal ones at scale factors 1 to 23 and all scale factors for the $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma^2}$, respectively. This demonstrates that the focal EEG signals are generally less complex than the non-focal ones, something that is in agreement with [38] and [39].

A student's *t*-test was also run for focal vs. non-focal signals. We adjusted the false discovery rate (FDR) [45] independently for each multivariate measure. The significance level of *p*-value tests was 0.05 for the EEG signals. The results demonstrate that the $RCmvMFE_{\sigma^2}$ approach achieves significant differences at all scale factors while the $RCmvMFE_{\mu}$ algorithm yields significant differences at scale factors between 1 and 20.

Using the $RCmvMFE_{\sigma^2}$ and $RCmvMFE_{\mu}$ values, we also employ a support vector machine (SVM) with linear kernel [46] to classify the focal and non-focal subjects in the WEKA data mining software [47]. We ran 100 repetitions of a 5-fold cross-validation applying at the subject level. The average classification accuracies were 96% and 100% for $RCmvMFE_{\mu}$ and

RCmvMFE$_{\sigma}^2$, respectively. The classification accuracy rates using mvMSE$_{\mu}$ and mvMSE$_{\sigma}^2$ were 92% and 100%, respectively. Although both the mvMSE$_{\sigma}^2$ and RCmvMFE$_{\sigma}^2$ lead to 100% accuracy rate, RCmvMFE$_{\mu}$ yields to higher accuracy in comparison with mvMSE$_{\mu}$. Note that the features were the values of entropy at each scale. To sum up, the refined composite-based techniques and/or variance-based coarse graining process has better performance to discriminate focal and no-focal signals.

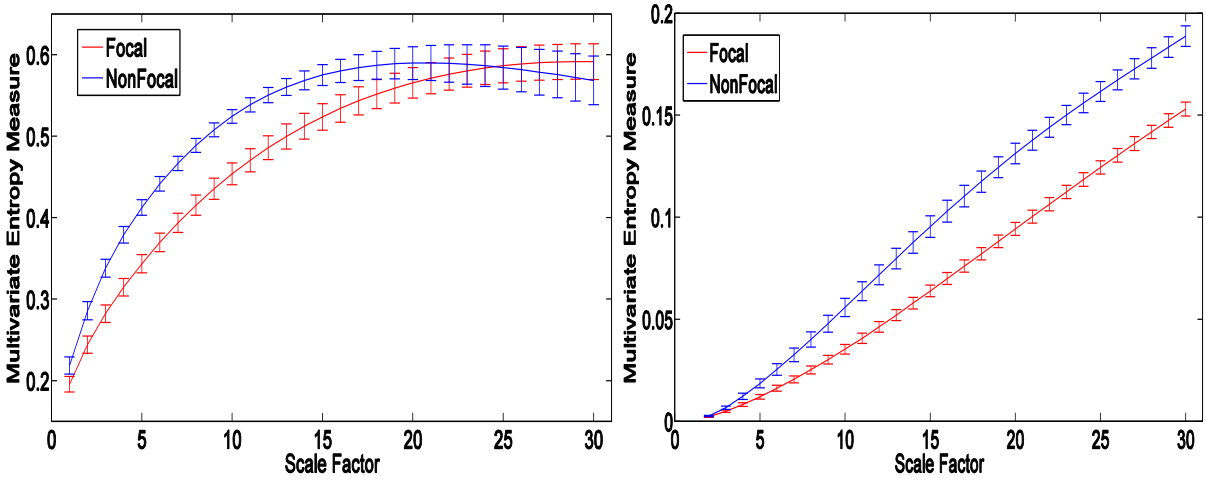
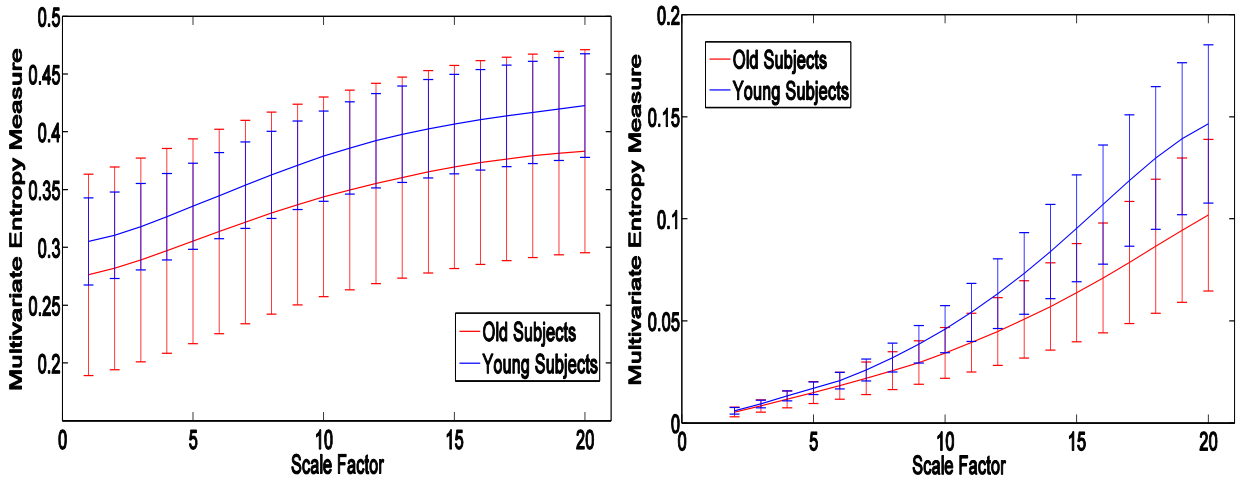


Figure 9. Error bars illustrating the mean ± SD of the (a) RCmvMFE$_{\mu}$ and (b) RCmvMFE$_{\sigma}^2$ values computed from focal and non-focal EEG signals.

We also apply the RCmvMFE$_{\mu}$ and RCmvMFE$_{\sigma}^2$ methods to the trivariate Fantasia time series to simultaneously analyze the complexity of continuous ECG, respiration, and blood pressure signals. The presence of long-range correlations in the ECG and respiration signals was earlier established using detrended fluctuation analysis [40, 48]. The univariate MSE method was applied separately to each of continuous electrocardiographic (ECG) and respiration signals [21].

The results showed that a lack of long-term correlations in both cardiac and respiratory dynamics, demonstrating the univariate MSE method was not able to produce robust estimates. Thus, we employ multivariate entropy-based methods. The error bars illustrating the distributions of the $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma}^2$ values computed from young and old subjects are shown in Figure 10. For each scale factor, the average of entropy values for young people are larger than that for elderly subjects using both $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma}^2$. For Fantasia database, our results are in agreement with those obtained by the other entropy-based approaches [21, 49, 50]. In addition, a *t*-test was performed for the young and old groups. The significance level of *p*-value tests was 0.05. We adjusted the FDR independently for each of $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma}^2$. The results show the $RCmvMFE_{\sigma}^2$ method achieves significant differences at scale factors 7 to 20 whereas the $RCmvMFE_{\mu}$ algorithm does not yield significant differences at scale factors between 1 and 20. This demonstrates a better performance of the $RCmvMFE_{\sigma}^2$ algorithm over the $RCmvMFE_{\mu}$ approach for the distinction of young and old subjects. We also applied the same classification scheme to discriminate the young and old groups. The average classification accuracies were 65% and 75% for $RCmvMFE_{\mu}$ and $RCmvMFE_{\sigma}^2$, respectively. This shows the importance of variance in coarse-graining process. The results obtained by the fuzzy membership function used in this paper are similar to those achieved by the function proposed in [15] but the proposed fuzzy membership function leads to a considerably faster method.



(a) (b)

Figure 10. Error bars illustrating the mean ± SD of the (a) RCmvMFE$_\mu$ and (b) RCmvMFE$_\sigma^2$ values computed from young and old subjects (Fantasia database).

In future work, we will investigate the suitability of RCmvMFE$_\sigma^2$ and RCmvMFE$_\mu$ methods to characterize other physiological time series to characterize Alzheimer's and Parkinson's diseases. Moreover, we intend to consider higher moments (e.g. the third moment or skewness) in the coarse-graining step of multivariate multiscale entropy-based methods.

5. Conclusions

This paper introduces RCmvMFE$_\sigma^2$ and RCmvMFE$_\mu$ to characterize the complexity of multivariate signals over multiple temporal scales that suits to consider the long-range within- and cross-channel correlations. These techniques are able to incorporate the simultaneous analysis of multichannel time series as a unique block within a multiscale framework. Extension of the MSE$_\sigma^2$ to multichannel time series has led to generalize the conventional mvMSE$_\mu$ to a family of statistics by using a different moment, i.e. variance, in the coarse-graining step. The

introduced fuzzy membership function has significantly decreased the running time in comparison with the existing $mvMFE_{\mu}$ while increasing the stability of the results. The proposed refined composite technique for multivariate time series has also led to improve the stability of the results for noise signals, especially in high temporal scales. The (RC)$mvMFE_{\sigma}^2$ and (RC)$mvMFE_{\mu}$ methods extract different kinds of dynamical properties (or features) of spread and mean, respectively, over multiple time scales. Based on this fact, which has been supported by the classification results, we conclude that both the (RC)$mvMFE_{\sigma}^2$ and (RC)$mvMFE_{\mu}$ offer complementary complexity profiles in signal analysis.

Appendix

The codes for our analysis, including mvSE, mvFE, $mvMSE_{\mu}$, $mvMFE_{\mu}$, $RCmvMSE_{\mu}$, $RCmvMFE_{\mu}$, $mvMSE_{\sigma}^2$, $mvMFE_{\sigma}^2$, $RCmvMSE_{\sigma}^2$, and $RCmvMFE_{\sigma}^2$ are available online at http://dx.doi.org/10.7488/ds/1432.

Acknowledgment

The authors would like to thank Eli Kinney-Lang from The Institute for Digital Communications, The University of Edinburgh, for helpful comments and suggestions. They extend thanks to Dr. Andrzejak from the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain, for providing the EEG data described in Section 3.

References:

[1] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences,* vol. 88, pp. 2297-2301, 1991.

[2] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology,* vol. 278, pp. H2039-H2049, 2000.
[3] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 15, pp. 266-272, 2007.
[4] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," *Physical Review Letters,* vol. 88, pp. 1-4, 2002.
[5] M. Rostaghi and H. Azami, "Dispersion Entropy: A Measure for Time Series Analysis," *IEEE Signal Processing Letters,* vol. 23, pp. 610-614, 2016.
[6] P. Micó, M. Mora, D. Cuesta-Frau, and M. Aboy, "Automatic segmentation of long-term ECG signals corrupted with broadband noise based on sample entropy," *Computer Methods and Programs in Biomedicine,* vol. 98, pp. 118-129, 2010.
[7] M. Ferrario, M. G. Signorini, G. Magenes, and S. Cerutti, "Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress," *Biomedical Engineering, IEEE Transactions on,* vol. 53, pp. 119-125, 2006.
[8] A. Holzinger, M. Hörtenhuber, C. Mayer, M. Bachler, S. Wassertheurer, A. Pinho*, et al.*, "On Entropy-Based Data Mining," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. vol. 8401, A. Holzinger and I. Jurisica, Eds., ed: Springer Berlin Heidelberg, 2014, pp. 209-226.
[9] H. Li, C. Peng, and D. Ye, "A study of sleep staging based on a sample entropy analysis of electroencephalogram," *Bio-Medical Materials and Engineering,* vol. 26, pp. 1149-1156, 2015.
[10] Y. Song and J. Zhang, "Discriminating preictal and interictal brain states in intracranial EEG by sample entropy and extreme learning machine," *Journal of neuroscience methods,* 2015.
[11] P. Luukka, "Feature selection using fuzzy entropy measures with similarity classifier," *Expert Systems with Applications,* vol. 38, pp. 4600-4607, 2011.
[12] M. Hu and H. Liang, "Intrinsic mode entropy based on multivariate empirical mode decomposition and its application to neural data analysis," *Cognitive neurodynamics,* vol. 5, pp. 277-284, 2011.
[13] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E,* vol. 71, p. 021906, 2005.
[14] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Physical Review Letters,* vol. 89, pp. 1-4, 2002.
[15] L. Peng, J. Lizhen, Y. Chang, L. Ke, L. Chengyu, and L. Changchun, "Coupling between short-term heart rate and diastolic period is reduced in heart failure patients as indicated by multivariate entropy analysis," in *Computing in Cardiology Conference (CinC), 2014*, 2014, pp. 97-100.
[16] L. E. V. Silva, B. C. T. Cabella, U. P. da Costa Neves, and L. O. M. Junior, "Multiscale entropy-based methods for heart rate variability complexity analysis," *Physica A: Statistical Mechanics and its Applications,* vol. 422, pp. 143-152, 2015.
[17] H. Fogedby, "On the phase space approach to complexity," *Journal of Statistical Physics,* vol. 69, pp. 411-425, 1992/10/01 1992.
[18] D. Bonchev and D. H. Rouvray, *Complexity: Introduction and fundamentals* vol. 7: CRC Press, 2003.
[19] A. L. Goldberger, C. K. Peng, and L. A. Lipsitz, "What is physiologic complexity and how does it change with aging and disease?," *Neurobiology of Aging,* vol. 23, pp. 23-26, 1// 2002.

[20] M. D. Costa and A. L. Goldberger, "Generalized Multiscale Entropy Analysis: Application to Quantifying the Complex Volatility of Human Heartbeat Time Series," *Entropy,* vol. 17, pp. 1197-1203, 2015.
[21] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy: A tool for complexity analysis of multichannel data," *Physical Review E,* vol. 84, p. 061918, 2011.
[22] D. Labate, F. La Foresta, G. Morabito, I. Palamara, and F. C. Morabito, "Entropic Measures of EEG Complexity in Alzheimer's Disease Through a Multivariate Multiscale Approach," *Sensors Journal, IEEE,* vol. 13, pp. 3284-3292, 2013.
[23] Z.-K. Gao, P.-C. Fang, M.-S. Ding, and N.-D. Jin, "Multivariate weighted complex network analysis for characterizing nonlinear dynamic behavior in two-phase flow," *Experimental Thermal and Fluid Science,* vol. 60, pp. 157-164, 2015.
[24] M. Ahmed, N. Rehman, D. Looney, T. Rutkowski, and D. Mandic, "Dynamical complexity of human responses: a multivariate data-adaptive framework," *Bulletin of the Polish Academy of Sciences: Technical Sciences,* vol. 60, pp. 433-445, 2012.
[25] F. C. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito, and I. Palamara, "Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG," *Entropy,* vol. 14, pp. 1186-1202, 2012.
[26] Z.-K. Gao, Y.-X. Yang, P.-C. Fang, Y. Zou, C.-Y. Xia, and M. Du, "Multiscale complex network for analyzing experimental multivariate time series," *EPL (Europhysics Letters),* vol. 109, p. 30005, 2015.
[27] Z.-K. Gao, M.-S. Ding, H. Geng, and N.-D. Jin, "Multivariate multiscale entropy analysis of horizontal oil–water two-phase flow," *Physica A: Statistical Mechanics and its Applications,* vol. 417, pp. 7-17, 2015.
[28] S.-D. Wu, C.-W. Wu, S.-G. Lin, K.-Y. Lee, and C.-K. Peng, "Analysis of complex time series using refined composite multiscale entropy," *Physics Letters A,* vol. 378, pp. 1369-1374, 2014.
[29] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy analysis," *Signal Processing Letters, IEEE,* vol. 19, pp. 91-94, 2012.
[30] L. Angelini, R. Maestri, D. Marinazzo, L. Nitti, M. Pellicoro, G. D. Pinna*, et al.*, "Multiscale analysis of short term heart beat interval, arterial blood pressure, and instantaneous lung volume time series," *Artificial intelligence in medicine,* vol. 41, pp. 237-250, 2007.
[31] S.-D. Wu, C.-W. Wu, K.-Y. Lee, and S.-G. Lin, "Modified multiscale entropy for short-term time series analysis," *Physica A: Statistical Mechanics and its Applications,* vol. 392, pp. 5865-5873, 2013.
[32] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series," *Physica D: Nonlinear Phenomena,* vol. 121, pp. 75-88, 1998.
[33] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical review A,* vol. 33, p. 1134, 1986.
[34] A.-M. Albano, A. Mees, G. De Guzman, and P. Rapp, *Data requirements for reliable estimation of correlation dimensions*: Springer, 1987.
[35] E. Sejdić and L. A. Lipsitz, "Necessity of noise in physiology and medicine," *Computer Methods and Programs in Biomedicine,* vol. 111, pp. 459-470, 2013.
[36] F. Diebold, *Elements of forecasting*: Cengage Learning, 2006.
[37] Y. Tang, W. Pei, K. Wang, Z. He, and Y. Cheung, "a Theoretical Analysis of Multiscale Entropy Under the Inverse Gaussian Distribution," *International Journal of Bifurcation and Chaos,* vol. 19, pp. 3161-3168, 2009.

[38] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Physical Review E,* vol. 86, p. 046206, 2012.
[39] R. Sharma, R. B. Pachori, and U. R. Acharya, "Application of Entropy Measures on Intrinsic Mode Functions for the Automated Identification of Focal Electroencephalogram Signals," *Entropy,* vol. 17, pp. 669-691, 2015.
[40] N. Iyengar, C. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology,* vol. 271, pp. R1078-R1084, 1996.
[41] D. Looney, A. Hemakom, and D. P. Mandic, "Intrinsic multi-scale analysis: a multi-variate empirical mode decomposition framework," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2015, p. 20140709.
[42] G. F. Reed, F. Lynn, and B. D. Meade, "Use of coefficient of variation in assessing variability of quantitative assays," *Clinical and diagnostic laboratory immunology,* vol. 9, pp. 1235-1239, 2002.
[43] H. Azami, K. Smith, A. Fernandez, and J. Escudero, "Evaluation of resting-state magnetoencephalogram complexity in Alzheimer's disease with multivariate multiscale permutation and sample entropies," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 7422-7425.
[44] S. M. Pincus and A. L. Goldberger, "Physiological time-series analysis: what does regularity quantify?," *American Journal of Physiology-Heart and Circulatory Physiology,* vol. 266, pp. H1643-H1656, 1994.
[45] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 289-300, 1995.
[46] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," ed, 2003.
[47] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD explorations newsletter,* vol. 11, pp. 10-18, 2009.
[48] C.-K. Peng, J. E. Mietus, Y. Liu, C. Lee, J. M. Hausdorff, H. E. Stanley*, et al.*, "Quantifying fractal dynamics of human respiration: age and gender effects," *Annals of biomedical engineering,* vol. 30, pp. 683-692, 2002.
[49] S. Nemati, B. A. Edwards, J. Lee, B. Pittman-Polletta, J. P. Butler, and A. Malhotra, "Respiration and heart rate complexity: effects of age and gender assessed by band-limited transfer entropy," *Respiratory physiology & neurobiology,* vol. 189, pp. 27-33, 2013.
[50] P. Li, C. Liu, K. Li, D. Zheng, C. Liu, and Y. Hou, "Assessing the complexity of short-term heartbeat interval series by distribution entropy," *Medical & biological engineering & computing,* vol. 53, pp. 77-87, 2015.